COGNITIVE NEUROPSYCHOLOGY, 2009, 26 (1), 1–6

Psychology Press
Taylor & Francis Group

Introduction

Letter recognition: From perception to representation

Matthew Finkbeiner and Max Coltheart
Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia

The development of written language is a remarkable cultural and cognitive achievement that is made all the more remarkable when one considers how recently in our evolutionary history it has developed. In fact, reading may represent the most complex skill that humans can master that cannot be attributed to a specific genetic predisposition. For cognitive scientists, reading constitutes an especially attractive domain of enquiry, since it involves a uniquely human and highly constrained form of visual pattern recognition. And letters are ideal as experimental stimuli for several reasons—they are highly overlearned visual patterns, they are simple to construct and easy to control, and, importantly, they are real, nameable objects.

What is it about reading that needs to be explained? While it would be nice to be able to explain how a reader, in the course of turning the pages of *The Brothers Karamazov*, comes to understand what it was like to live in Imperial Russia in the nineteenth century, the goal of the work reported here is a little more modest: the explanation of how a letter is recognized. Understanding letter recognition is a part—even though only a very small part—of understanding what happens when a novel is read. At the very minimum, success in reading novels involves determining the location and identity of the letters that comprise a written word. We can't have a complete understanding of reading if we don't understand letter recognition.

Since the time of Dejerine (1892), it has been well accepted that visual word recognition proceeds via a "visual word form" representation. Interestingly, though, the majority of the contemporary work on visual word recognition begins with the word form, thereby glossing over the steps that are required to go from the printed letter string on the page to the visual word form stored in long-term memory. One goal of this special issue is to bring together both functional and neural perspectives on the earliest stage of the reading process—namely, *letter perception*.

Letter perception involves the assembly of visual features into letters. Somewhat surprisingly, this front end of the reading system has largely been overlooked in the word recognition research. While the interactive activation (IA) model (McClelland & Rumelhart, 1981) and its derivatives (e.g., dual route cascaded, DRC, model, Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; multiple readout model, MROM, Grainger & Jacobs, 1996) include a feature level that feeds into a subsequent letter level, the architectural and dynamical assumptions underlying these first two levels of processing have largely gone untested. This is surprising

Correspondence should be addressed to Matthew Finkbeiner, Macquarie Centre for Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia (E-mail: matthew.finkbeiner@maccs.mq.edu.au).

http://www.psypress.com/cogneuropsychology
DOI:10.1080/02643290902905294

because determining how readers extract letter identities from the highly variable featural information in the input is fundamental to attempts to understand the reading process. The first three articles in this special issue focus on letter perception—that is, the way in which we construct letters from visual features.

In the opening article, Rey and colleagues introduce an innovative and promising way to use neurophysiological data to test and constrain cognitive models. They report an event-related potential (ERP) study in which they first seek to establish the time course of letter-specific processing. Using ERPs, they demonstrate that the brain response to letters versus nonletters begins to diverge at approximately 150 ms from stimulus onset. They then use the results of their ERP study (latency and amplitude data) to test different instantiations of the IA model of word recognition (McClelland & Rumelhart, 1981). Interestingly, Rey et al. find that their ERP results are consistent with a version of the IA model in which feature-to-letter level connections are excitatory but not inhibitory and in which there is lateral inhibition between letter units as well as letter-to-feature feedback. Versions of the model that included inhibitory feature-to-letter connections were found to be inconsistent with their results.

Suppose in the IA model and in the models derived from it (MROM, DRC), we did set the parameters in the way Rey and colleagues conclude that we should: What might the consequences for the models be? Table 1 shows the relevant parameters used by the DRC model (inherited from the IA model, which uses the same values for these parameters) and contrasts these with the recommendations of Rey and colleagues.

It will be an interesting exercise in computational modelling to see whether, if these parameter changes are made to the DRC model, this compromises the model's present ability to simulate a wide variety of results obtained in human reading-aloud and lexical-decision studies. We won't embark on that exercise here, but will make some remarks about this.

First, some DRC simulation work currently in progress with Claudio Mulatti has led us to the view that letter-to-letter inhibition should not be zero, and that is one of the recommendations of Rey and colleagues. This work has to do with the variable known as letter confusability. Any pair of letters has a visual similarity that can be defined by, for example, the number of features the two letters have in common. The greater this similarity is, the more likely readers will be to confuse the two letters; so letter similarity is sometimes referred to as letter confusability. For any letter, one can average its confusabilities with the 25 other letters to give a measure of that particular letter's overall letter confusability (LC).

Table 1. *Parameters at the feature and letter levels for the DRC model and as recommended by Rey and colleagues*

	DRC model	*Rey et al. recommendations*
Feature-to-letter excitation	.005	nonzero (same as DRC)
Feature-to-letter inhibition	.150	.000
Letter-to-feature excitation	.000	nonzero
Letter-to-feature inhibition	.000	.000 (same as DRC)
Letter-to-letter lateral inhibition	.000	nonzero

Note: DRC = dual route cascaded.

And for any letter string, one can measure its overall confusability, either as TLC (total letter confusability—the sum of the confusabilities of the letters in the string) or as MLC (mean letter confusability—the mean of the confusabilities of the letters in the string).

Gosselin, Blais, and Arguin (2006) showed that the characteristic symptom of patients with pure alexia, a highly exaggerated effect of number of letters on reading latency for words, is absent if the TLC of letter strings is matched across word lengths. Thus, it appears that effects of this kind previously reported in pure alexics arise because these patients are abnormally sensitive to letter confusability. This offers the prospect of DRC simulation of this form of acquired disorder provided that the DRC model's reading is sensitive to TLC. Now, if an appreciable proportion of the features characterizing some letter—say the letter E—are shared by some other letter—say F—then presentation of an E will partly excite the letter representation for F, exemplifying the confusability of F and E. But this won't harm the correct letter E, because the lateral inhibition between F and E is set to zero, and so the activation level of E won't be affected by F being somewhat activated. However, if between-letter lateral inhibition is set to some positive value, one can now see effects of TLC on the DRC model's reading latencies, and we have observed circumstances with human readers where their reading is affected by TLC too. To capture these effects (and perhaps ultimately to simulate pure alexia) we have had in our current project to adopt a nonzero setting for letter-to-letter lateral inhibition (all published work with the IA model and its descendants have used a zero value for this parameter). Rey and colleagues have offered an independent reason for doing this.

Secondly, setting feature-to-letter inhibition from a positive value to zero and setting letter-to-feature excitation from zero to a positive value will produce much more rapid and extensive activations of letters, including incorrect letters, than occurs with the standard parameters. A visual feature present in the input won't inhibit letters that don't possess that feature, and positive feedback from letters to features will in turn generate activations for incorrect letters that share features with the presented letter. It remains to be seen whether accurate letter recognition, and accurate word recognition, is achievable under these conditions. Nonword reading might also be threatened by lexical capture here. Only a programme of simulations will tell us whether or not there are problems with these two parameter changes recommended by Rey and colleagues.

In the second article, Fiset and colleagues use the *Bubbles* technique to try to learn something more about what the features from which letter representations are constructed actually are. Although the IA model did have an explicit and functional visual feature level (it had to, otherwise it could not have been a runnable computational model), that was not meant to be a realistic characterization of the feature set that human readers use for letter recognition. The IA model and its descendants use a 14-feature system devised by Rumelhart and Siple (1974) and shown in Figure 1.

This means not only that these models cannot recognize lower-case letters but also that upper-case letter recognition is restricted by font: The models can recognize A but not **A**. There are still no computational models of reading that have progressed beyond this extremely crude way of representing the visual feature level of the reading system.

The results of Fiset and colleagues reveal that line terminations and horizontals are the most important letter features in letter identification. Additionally, they show very nicely that different letter features follow different time courses in letter identification. That is, not only are some features more important than others in the identification of letters, but it is also the case that individuals extract particular features before they do others. This is a very interesting finding that is sure to motivate a new line of investigation comparing letter and nonletter objects. It also opens up a way forward for computational models of reading to incorporate a somewhat more psychologically realistic visual-feature level, guided by the findings of Fiset and colleagues. However, although this work has provided more sophisticated ideas about what kinds of things might constitute the visual features of letters, it still has not gone beyond a particular font and a particular case.

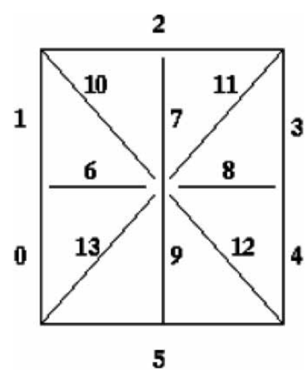

Figure 1. *The Rumelhart–Siple feature set. The letter A, for example, would be represented by the binary feature pattern 11111010100000.*

Although each of their participants saw a letter on 26,000 trials, it was always an uppercase Arial letter. We don't know whether the particular letter features identified in the study, and the time courses of their importance, would be the same if the stimuli had been upper-case Courier letters—or even lower-case Arial letters.

In the third article, Pelli and colleagues test the Gestalt claim that object recognition, in this case letter identification, depends upon the grouping of features. These authors created letter-shape contours and displayed them in visual noise. They then perturbed the collinearity of the segments that comprised the letter shapes by rotating, offsetting, or phase-shifting the successive segments. Not surprisingly, Pelli et al. found that observers' accuracy in an identification task decreased as the severity of the perturbations increased. Importantly, this inverse relationship held equally for all three types of perturbations. This finding demonstrates that the effect of good continuation on letter identification is independent of the specific type of perturbation, which, in turn, confirms the importance of grouping in letter identification.

Moving on from letter perception, the second goal of this special issue is to focus on the nature of the letter representations themselves, and especially upon the concept of abstract letter representation (and how children acquire such representations in the course of learning to read). An abstract letter representation (Coltheart, 1981) is abstract in the sense that it is independent of font and case. At the level of abstract letter representations, an **a** is an a, and so is an A and an **A**. At this level of the reading system, there is just one unit (an abstract letter unit or ALU) representing the first letter of the alphabet, and it is supposed to be activated regardless of the font, case, or style in which that letter appears on the page. The same of course is true for all of the other letters of the alphabet: Each has its own ALU.

There are various lines of evidence consistent with the hypothesis that the human reading system includes an ALU level. First, it has been demonstrated using a variety of experimental paradigms that when words or letters are presented very briefly participants are sometimes able to report what letter or word they had seen without being able to report whether it was in upper or lower case (Adams, 1979; Coltheart & Freeman, 1974; Friedman, 1980; McClelland, 1976). Secondly, some brain-damaged patients who have lost all access to the names of letters can nevertheless perform cross-case matching (deciding which of **a** or **e** matches A) perfectly (see e.g., Mycroft, Hanley, & Kay, 2002): If the patients aren't performing this task by matching letter names, it is hard to see how else they might be doing it except by matching abstract letter identities. Thirdly, Kinoshita and Kaplan (2008) reported that cross-case priming in a letter-match task is no greater for letters that are very similar across case (c/C) than for those that are not (a/A).

In the fourth article, Thompson reviews a large body of work on the development of ALUs in children as they learn to read. His review concludes that the evidence of ALU development in children is consistent with the common context hypothesis first proposed by Polk and Farah (1997). According to this hypothesis, children learn that "A" and "a"'are two variants of the same letter by encountering words like "cap" and "CAP". In this example, the first and last letters of the word do not vary much between upper and lower case, but the middle letter (a/A) does. On this hypothesis, the child is able to conclude that "a" is the same as "A" by virtue of the "common context" provided by the letters that do not vary between upper and lower case.

While Thompson concludes that the extant evidence is largely consistent with the common context hypothesis, he notes shortcomings and indicates that training studies are needed to test specific claims of the common context hypothesis. The next article by Polk and colleagues does just this. In their article, Polk and colleagues first demonstrate empirically that arbitrary symbols presented in common contexts during a training phase become more confusable with one another at test. They then extend the original Polk and Farah (1997) model to be used with real letters as input and successfully simulate their experimental findings.

Here again, though, current computational models of the reading system are seriously deficient. They posit that letter recognition involves an initial stage of visual feature analysis and a final ALU stage. Recognition of a letter is achieved when an appropriate ALU is sufficiently activated by the activated visual features, and communication between the two levels is achieved via excitatory or inhibitory links between particular visual feature units and particular ALUs. Thus in the IA model, the visual-feature unit for the feature *vertical line at left* would have an excitatory connection to the ALU for the letter **D**, and the visual-feature unit for the feature *vertical line at right* would have an inhibitory connection to the ALU for the letter **D**. What, then, if the input is the letter **d**? With these feature connections, a **d** would not excite the ALU for **D**, as it ought to: Indeed, it would inhibit its own ALU. No current computational model of reading addresses this issue.

How might the issue be addressed? One possibility is for there to be a level of case-specific letter units (CSUs) interposed between the visual feature level and the ALU level. So **D** and **d** would have distinct CSUs, but both of these CSUs would have excitatory links to the ALU for **D**/**d**. Petit, Midgley, Holcomb, and Grainger (2006) have discussed such ideas, though no one has yet attempted to implement them computationally. Of course, font and style variation will be a problem here. If **a** and ɑ have a single common CSU, how are the features to be hooked up to this CSU so that both **a** and a will excite and not inhibit the **a**/ɑ CSU?

Finally, we conclude with two articles that use functional magnetic resonance imaging (fMRI) to look at how patterns of neural activity evoked in the left ventral visual pathway by nonletters come, over time, to mimic the patterns of activity that are evoked by letters. In the first of these articles, James and Atwood had participants learn to recognize pseudoletters in three different training paradigms: by writing, by typing, or by visual practice only. Very interestingly, they report that neural activation in the left fusiform was greater for the pseudoletters after training than it was for similar but untrained stimuli, but only for those individuals who had learned the pseudoletters by writing. The authors take these findings to suggest that the functional organization of the visual system is "shaped" in a nontrivial way by the motor systems that are recruited when interacting with particular stimuli. While the interesting implications of this study extend beyond the specific way in which individuals learn to recognize letters, they also provide support for the longstanding idea of a multisensory approach to the treatment of reading difficulties (Hulme, 1981), not to mention the slogan of the Spalding reading treatment programme ("The writing road to reading": Spalding & North, 2003).

In the final article, Wong and colleagues report an fMRI study that investigated the roles of expertise and stimulus properties in the neural representation of Roman letters and Chinese characters. They used a within-subjects design with Chinese–English bilingual readers and English (non-Chinese) readers. Using objects as a baseline, they identified regions of cortex that responded selectively to Chinese characters (in the bilinguals), Roman characters in the non-Chinese readers, and faces (for both groups). They then investigated the degree to which the Roman- and Chinese-selective cortical areas overlapped. Interestingly, they found that these regions overlapped almost perfectly in each of the bilingual participants. In contrast, non-Chinese exhibited much more variability in their responses, with some not responding at all to Chinese characters (relative to objects), but with others responding similarly for both letters and characters. Taken together, these results implicate the roles of both expertise and stimulus properties in the functional specialization of the ventral visual cortex.

In conclusion, the seven articles in this special issue tackle the earliest stages of the reading process. The first three articles address issues of letter perception—that is, how letters are extracted from visual features. The remaining four articles address the nature of letter representations from both functional and neural perspectives. These articles introduce novel and interesting ways to investigate the very earliest stages of the reading process. In doing so, we expect that the research reported here will generate future investigation in this highly tractable, yet long overlooked, area of reading research.

REFERENCES

Adams, M. J. (1979). Models of word recognition. *Cognitive Psychology*, *11*, 133–176.

Coltheart, M. (1981). Disorders of reading and their implication for models of normal reading. *Visible Language*, *3*, 245–286.

Coltheart, M., & Freeman, R. (1974). Case alternation impairs word identification. *Bulletin of the Psychonomic Society*, *3*, 102–104.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Dejerine, J. (1892). Contribution à l'étude anatomo-pathologique et clinique des différentes variétés de cécité verbale. *Mémoires de la Société de Biologie*, *4*, 61–90.

Friedman, R. B. (1980). Identity without form: Abstract representations of letters. *Perception and Psychophysics*, *28*, 53–60.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple readout model. *Psychological Review*, *103*, 518–565.

Hulme, C. (1981). *Reading retardation and multisensory teaching*. London: Routledge and Kegan Paul.

Kinoshita, S., & Kaplan, L. (2008). Priming of abstract letter identities in the letter match task. *Quarterly Journal of Experimental Psychology*, *61*, 1873–1885.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 80–91.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375–407.

Mycroft, R., Hanley, J. R., & Kay, J. (2002). Preserved access to abstract letter identities despite abolished letter naming in a case of pure alexia. *Journal of Neurolinguistics*, *15*, 99–108.

Petit, J. P., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2006). On the time course of letter perception: A masked priming ERP investigation. *Psychonomic Bulletin and Review*, *13*, 674–681.

Polk, T. A., & Farah, M. (1997). A simple common contexts explanation for the development of abstract letter identities. *Neural Computation*, *9*, 1275–1287.

Rumelhart, D. E., & Siple, P. (1974). The process of recognizing tachistoscopically presented words. *Psychological Review*, *81*, 99–118.

Spalding, R. B., & North, M. E. (2003). *The writing road to reading: The Spalding method for teaching speech, spelling, writing, and reading.* New York: HarperCollins.